PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	COVID-19 International Border Surveillance Cohort Study at Toronto's Pearson Airport
AUTHORS	Goel, Vivek; Bulir, David; De Prophetis, Eric; Jamil, Munaza; Rosella, Laura; Mertz, Dominik; Regehr, Cheryl; Smieja, Marek

VERSION 1 – REVIEW

REVIEWER	Migisha, Richard Mbarara University of Science and Technology
REVIEW RETURNED	19-Apr-2021

GENERAL COMMENTS	General comment This study assessed positivity rates for SARS-CoV-2 at an International Airport in Canada at different days; findings from the study inform cost-effective duration of quarantine in various countries. Overall, this is a well-executed study, and limitations are well acknowledged. However, I have some comments: 1. Abstract: Line 25 in the results section, it states, “Of 16,361 passengers enrolled, 248 (1·5%, 95% CI 1.3%,1.5%) tested positive”. The authors should revisit this since the point estimate is not included in the confidence interval. Please check the upper limit value and revise accordingly. The same applies to line 9 under results section, sub-section Symptoms and Positive COVID-19 Results. 2. The authors assessed the mental health of the participants and found the deterioration of the mental status with longer duration of stay. However, the scoring system used here is not very clear in the methods. Please clarify this. 3. Under discussion, the authors state, “Preliminary analysis suggests that males reported being less likely to comply with public health recommendations (data not presented).” I would suggest the authors limit their discussion to the presented results to avoid speculations. This should be revised. 4. Under discussion lines 88-42, the authors hint on the findings of Russell et al in this paragraph. However, it is not clear how these results re-enforce or disagree with theirs. This needs to be revised by giving more clarity.

REVIEWER	Cook, Alex R National University Singapore Saw Swee Hock School of Public Health
REVIEW RETURNED	27-Apr-2021

GENERAL COMMENTS	The authors report an observational participatory study in which travellers arriving into Toronto were to swab themselves once a week over two weeks. The results are not surprising, as they found more tested positive on arrival than at t=7 or t=14. Despite the lack of surprise, the study adds information that may be valuable in formulating border policies. I have some minor suggestions: Abstract: The authors claim that "A single arrival test will pick up two-thirds of individuals who will become positive, with most of the rest detected on the second test at day 7." Technically, this is only on condition that tests are done at {0,7,14}. In the universe in which tests can be done after day 14, this estimate should decline (by a little). 5.13: Sentence is illogical. If the parenthetical example is omitted the sentence becomes: "Many countries… have kept borders closed to foreign travelers, with the exception of essential workers and returning Canadians" 5.14: it doesn't "minimise" the risk: that would involve complete closure of the border. "reduce" would be more correct. 5.51: consider replacing 'fall' by the months or quarter, since it's an international journal and not all countries experience the same seasons and at the same times. 7.3: Even unweighted estimates don't follow a binomial distribution…! Asymptomatic normal distribution for the MLE, or a beta posterior, surely. 9.9: Standardise decimal place symbols here and throughout 9.9: Please check CI upper bound. 9.15: Confirm that the 3 who were symptomatic on arrival were not excluded as per the criteria in the methods section, and explain the discrepancy. 10.33: It strikes me as odd that you would recommend imposing the same requirements on travellers from India and New Zealand. F4: What is the point of the figure? There are only three numbers to report, you could just write them down.

VERSION 1 – AUTHOR RESPONSE

Reviewer: 1

Dr. Richard Migisha, Mbarara University of Science and Technology

1. Abstract: Line 25 in the results section, it states, "Of 16,361 passengers enrolled, 248 (1·5%, 95% CI 1.3%,1.5%) tested positive". The authors should revisit this since the point estimate is not included in the confidence interval. Please check the upper limit value and revise accordingly. The same applies to line 9 under results section, sub-section Symptoms and Positive COVID-19 Results.

This has been corrected to 1.7% in the abstract and results.

2. The authors assessed the mental health of the participants and found the deterioration of the mental status with longer duration of stay. However, the scoring system used here is not very clear in the methods. Please clarify this.

An expanded clarification on how well-being/ mental health was measured has been added to Methods: Data management and analysis. Note, we used the WHO-5 Wellbeing Index, a validated tool which is recommended for use in the context of the pandemic by the WHO.[1]

3. Under discussion, the authors state, "Preliminary analysis suggests that males reported being less likely to comply with public health recommendations (data not presented)." I would suggest the authors limit their discussion to the presented results to avoid speculations. This should be revised.

This statement has been deleted.

4. Under discussion lines 88-42, the authors hint on the findings of Russell et al in this paragraph. However, it is not clear how these results re-enforce or disagree with theirs. This needs to be revised by giving more clarity.

We have revised to provide clarity on the uses and limitations of our results, and the complementary value of models such as those developed by Russell et al.

Reviewer: 2

Dr. Alex R Cook, National University Singapore Saw Swee Hock School of Public Health

Abstract: The authors claim that "A single arrival test will pick up two-thirds of individuals who will become positive, with most of the rest detected on the second test at day 7." Technically, this is only on condition that tests are done at {0,7,14}. In the universe in which tests can be done after day 14, this estimate should decline (by a little).

The reviewer is correct that there is the possibility of individuals becoming positive after day 14. While there are isolated reports of cases after 14 days, the expected number will be very small. We have revised the abstract to say "A single arrival test will pick up two-thirds of individuals who will become positive by day 14". A statement and reference to support this have been added to study limitations.

5.13: Sentence is illogical. If the parenthetical example is omitted the sentence becomes: "Many countries… have kept borders closed to foreign travelers, with the exception of essential workers and returning Canadians"

Agreed. Changed "Canadians" to "citizens and permanent residents."

5.14: it doesn't "minimise" the risk: that would involve complete closure of the border. "reduce" would be more correct.

Agreed. Done.

5.51: consider replacing 'fall' by the months or quarter, since it's an international journal and not all countries experience the same seasons and at the same times.

[1] World Health Organization. Regional Office for Europe. Survey tool and guidance: rapid, simple, flexible behavioural insights on COVID-19: 29 July 2020 [Internet]. Copenhagen; 2020. Available from: https://apps.who.int/iris/handle/10665/333549

Agreed. "Fall" changed to "September and October".

7.3: Even unweighted estimates don't follow a binomial distribution…! Asymptomatic normal distribution for the MLE, or a beta posterior, surely.

At each time point, an individual can test positive (i.e., success in a Bernoulli sense) or negative (i.e, failure). We present the actual number of people who tested positive at each time point. Unweighted tables are reporting the ***count*** *of cases by each timepoint with a 95% CI around the counts. As a binomial distribution is used for count data (i.e. positive vs negative result for COVID-19), using the binomial exact method is the most fitting choice. The rates are then calculated by dividing these counts by the number of participants registering a test at each timepoint. We do not have any reason to believe the proportions would follow a normal distribution; however, if we did assume the normal distribution the confidence intervals would be narrower, or less conservative, than what we have presented. The weighted tables used the bootstrap method due to the complexity introduced by weighting the number of experiments (i.e. tests) by the weights generated to account for loss to follow-up.*

We have clarified in Methods: Data Management and Analysis the distinction between the counts and rates.

9.9: Standardise decimal place symbols here and throughout

Agreed, done.

9.9: Please check CI upper bound.

Thank you for identifying this. It has been corrected to 1.7%.

9.15: Confirm that the 3 who were symptomatic on arrival were not excluded as per the criteria in the methods section, and explain the discrepancy.

Given the possible time lag between submitting a test on arrival and answering the online questionnaire (i.e. could have been later that day) it is possible that these 3 individuals were asymptomatic at the time they took the test, but became symptomatic later. They were not excluded from the case count. This has been clarified in Results: Symptoms and Positive COVID-19 Results. (All participants had to have reported no symptoms to the border officer in order to be eligible for the study, since anyone reporting symptoms on arrival was immediately referred to a quarantine officer.)

10.33: It strikes me as odd that you would recommend imposing the same requirements on travellers from India and New Zealand.

We make an observation on the difficulty of imposing country specific requirements based on epidemiological data, since circumstances change rapidly. However, this does not necessarily translate into our recommending the same requirements for all countries, the Russell reference does outline some of the factors that could be considered in developing such recommendations. The main purpose of our study is to support optimum testing and quarantine interval policies, additional inputs are required for determining when and to which regions such policies should be applied.

F4: What is the point of the figure? There are only three numbers to report, you could just write them down.

The figure was intended to emphasize the considerable drop in average WHO-5 score through the course of quarantine. We have replaced the figure with a sentence in results.

VERSION 2 – REVIEW

REVIEWER	Migisha, Richard Mbarara University of Science and Technology
REVIEW RETURNED	10-Jun-2021

GENERAL COMMENTS	The comments were well addressed and the quality of the paper is enhanced. I have no other comments.

REVIEWER	Cook, Alex R National University Singapore Saw Swee Hock School of Public Health
REVIEW RETURNED	06-Jun-2021

GENERAL COMMENTS	Thank you for addressing my previous comments.